

A)

	Motif 1	Motif 2	Motif 3	Motif 4
Rv2837c_Mt	VGWVCHVPEADDTIGAGLALAVLDGCG [35]	VDLVVTVIPSDVRLGALG [6]	RELLVDDHFAAND [10]	SADSTTTWVAEILDANGKPIDPRVAHCIYAGLATDTGSRFASV
s111253_Ssp	DLILCHQTDFEDVLGAAGVLAHLHFGSR [34]	IRSLYIVNQOQDRGLKAA [8]	ROVAIVDHEINSP [11]	AVGASTTLIVEKQORADISLSMVEASVMALGIHVDTGSLTFTQT
MGPA_Mg	IVIFHHVRPDCGLGAQQGLFHLIKANF [32]	EALAIWDAANYKNRIELRE [7]	KAVLRIDHEPNED [11]	SYVACCEQIVEMATVAKWTIEPVAATLLYIGIYTDSSNRFYSNT
YTQI_Bs	IILHRHVRPDCAYGQCGLTEILRETY [30]	GALVIVQDTANQERIDDOQR [4]	AKLMKIDHEPNED [13]	SVSEMIYELYLEGKEHGKLNKAAELIYAGIVGDTGRFLFPNT
AF2029_Af	LGIFTHDNEPDSMSSAYALRELAKQPD [37]	YDAFAIVDSGGPGVNSIP [3]	DISIVVDHHPAEK [10]	DVGATATILTEYIKELKITPSKILATALLFFGIKSETDEFKRNR
MJ0977_Mj	NKILIVTHIDTGLTSRAILQKLAERLN [23]	YDLIIPADLGSGQLKMIKE [11]	DKIIIIDHHEQPEE [18]	GAEICGAGVSYLEFAKAINNDWIDLAKYAVLGAAGDIQNIIEGKLI
MJ1198_Mj	RPIIIRHHEDTGYCGGIALEKAILPII [46]	LPLIVLIDNGSTDEDIPAI [7]	IEVIVDHEFPGE [85]	KORTYDREYLEKIALCMDFEAFYLRFMDGKGIVDDILATNIKEF
HP1042_Hp	MQVYHLSHIDLDGYACQLVSKQFFKNIQ [27]	EFLILVSDNLNLNEAEYL [13]	IQIQLDHEFISGK [19]	IVIEFLKKHYAILEPKNTTWLEPLVEMVNSVDINDTQGYGFEIG
RecJ_Hi	QKIVIVGDFDADGATSTALSVLARQLG [31]	VQLLMTVDNGVSSFDGVAF [5]	IRVLVDHELPPE [33]	LAVRAKPRELGIFTAETQPNFTDLDLVALGTIADVVPLDQNNR
RecJ_Hp	TEILVVGDDYDADGVISSAIMAKFFESLN [27]	APLIITVDNGINAFEAARF [5]	YTLIITDHECLHH [27]	LVAFYLCYGIHQLLGKEKSHSELCLAGVATIAQMMPLTFFNR
RecJ_Ssp	EKVITNGDADAGGITSTAVLWEGLGQFF [32]	TKLIVTQDTGSTNLDEIVY [5]	MDVIVDHEHTLPD [27]	VAFKLVREALYNQYPTVPQQPLEDLDDLVAIGLIADLVLTLOGDCR
YYBQ_Bs	ILIFGHQNPDDTICSIAIAYADLKNKLG [36]	VNGVILVDHNERQQSIKDI [3]	QVLEVDHERIAN [12]	PVGCTATILNKMYKENNVKIEKEIAGLMLSATISDSLLFKSPIC
ICRA_Sg	ILVFCHQNPDSDAIGSSYAFAYLAREAY [38]	AEQVILTCHNEFQQSVADI [3]	EVYGVVDHERVAN [12]	PVGSASSIVYRMFKESVAVSKEIAGLMLSGLISDTLLLLKSPTT
MJ0608_Mj	RYVVGHNKPDTSIASAIVLAVFLDCVP [31]	GKEIILVDHSEKSSQSFDDL [3]	KLIAVDHEKUGL [19]	IAELYFKDAIDLIGGKKKELKPDLAGLLLSAISDTVLFKSPIT
AF0756_Af	VYVVGHNKPDTSVCSAIAFAYLVNKNK [46]	GKKVALVDHSEKAQTVDGI [3]	EVVATVDHEKIGD [12]	PVGCTATVIKLLFDKTGVEIPKDIAGILLSSILSDTVIFKSATT
U60409_Lm	TVVQNEGQDMDSIVGCIYLAFLDKQF [44]	QIAHNLVDIAALNASVVLY [14]	RVVGVDHEFDEQ [11]	LRTVGSACTLVTELYREGGEDVVCPTLLTAPIVLDTVNFEPAQK
PPX1_SC	TICVGNESPDMDSIASAITYSYCYIYN [62]	ELNSYLVNDNDTPKNLKNY [2]	NVVGIDHHEFDLQ [14]	SCSSLVFNWYVEKLGQDREVVMNIAPLLMGAILIDTNSMRRKVE
PRUNE_Dm	HLVMGNESODLSAVSAVTLAFVYAQRH [48]	DUNVILVDHHVUSPLAPNVT [1]	NVTEIIDHRPLED [18]	SVGSCATLVAQRYLAEDQPRSTSVAAQLLHATIVLDTINFAPAAK
h-prune.1_Hs	HVVLGNEAOQLDSTVSALALAFYLAKTT [52]	QETLILVDHHILSKSDTAL [2]	AAVEVIDHRPIEP [12]	LVGSCATLVTERILQGAPEILDRTAALLHGTIILDCCVNMDLKIG

	Specific motifs	GIS
Rv2837c_Mt	[107] TVNLAAVASGFGGGHRLAAGYTTTGS	1648883
s111253_Ssp	[123] DTDLTQLLEPYGGGCHQAQAAVNLRDV	1653244
MGPA_Mg	[102] GINVRDIAIKYGGGGHNNASGAILTNK	1045875
YTQI_Bs	[103] GPVINGLARKYNGGGHPLASCASIYSW	2293259
AF2029_Af	[102] EVLRAFCDVGSAGGHAHAAGAQLPLG	2648507
MJ0977_Mj	[254] AIKYASEKVNCGGGHKFACGAYIPDN	2128614
MJ1198_Mj	[109] QLMEEIPEASLDGGGHECAGSLKFVEG	2128720
HP1042_Hp	[137] CDVCELSQCMFCNGGGHNRNASGGKIDGF	2314198
RecJ_Hi	[168] RIHSQHPNMILKFGGHMAAGLSIREE	1172895
RecJ_Hp	[155] DALNGVSSLLLYGGGHRQACGLSVEKN	2313437
RecJ_Ssp	[165] ALLHSQRHMLGFGGHPFAAGLSLPLD	1652638
YYBQ_Bs	[31] DLSKKTVEELISLDAKEFTLG [75]	TALLKGVSRRKKQVWPVLTD 586817
ICRA_Sg	[31] NLASKSAEELIDIDAKTFELN [75]	HAFLAGAVSRKKQVWPQLTE 1743856
MJ0608_Mj	[32] VVGKLKPEEIIINMDFKNFDFN [74]	SVFLEGVMSRRKKQVWPPLER 1591318
AF0756_Af	[32] AVDDLTAAMDIIKRDYKXDFDMS [75]	SVNLGVMSSRRKKQVWPPLER 11498362
U60409_Lm	[36] DVLALSVPQILRRDYKQFSFK [105]	YSLSDPSISRRK-LVPAISE 1407725
PPX1_SC	[41] DIKGFVSVDILKDYKQFNQ [97]	MPKQLNVEATRRKKQVWPYLEE 730369
PRUNE_Dm	[37] DISKLTLEVLKDKMKVLQTD [97]	LRQHNVOATRRK-ILPIVKKR 1079081
h-prune.1_Hs	[34] DVSGLTTEQMLRKDKQRTIYRQ [94]	YLOQNTQVSRKK-LLELLQEE 24308263

Family 1

Family 2

Fig. 1